Supplimentary material for "Research at the Auction Block: Problems for the Fair Benefits Approach to International Research"

Alex John London
Carnegie Mellon University

Kevin J.S. Zollman
Carnegie Mellon University

Appendix A Underlying Assumptions

How plausible are the assumptions that are required in order for the revenue equivalence theorem to hold? First, let the various costs to host communities from hosting a research initiative be represented as c_i, where i represents each potential host community. From the point of view of the other communities this value is a random variable independently drawn for each country from some distribution with a cumulative distribution $p(\cdot)$. Let v represent the surplus value generated by the study for the researcher. We then assume:

- **Supply and Demand Imbalance**: At any time, the number of potential host communities is greater than the number of potential research studies by at least one.
- **No Risk Seeking**: Parties to a bargain do not seek risk for its own sake.
- **Highest bidder**: Whatever the price paid, researchers will choose the community that represents the lowest cost to them.
- **Last place**: If a community has the highest possible cost, it expects to get zero surplus from the negotiation process.
- **Smooth distribution**: $p(\cdot)$ is strictly increasing and atomless over some range $[c_{min}, c_{max}]$

- **Common knowledge**: The value of the research, the distribution of costs, $p(\cdot)$, the structure of the bidding process, and the number of potential bidders is all common knowledge.

These assumptions appear to present a fair representation of the situation of most LMIC communities. The supply and demand imbalance represents the fact that for most diseases there are more communities that could host a research project that targets a given disease than there are studies targeting that condition. This does not require that there be more communities than research projects in total since a single community can host multiple studies. In fact, all that is required is that there are at least two potential host communities for each study.

The no risk seeking assumption is very weak. It merely requires that no party to one of these bargains has a positive utility for risk, as such. This does not mean that the various communities will not differ in their tolerance for risk. It means that they are not like mountain climbers or sky-divers, where they seek out this activity precisely because it is risky. This assumption does not prohibit risk aversion, but we do not assume it. If we did, then the expected profit for the researcher would actually increase since a community with a low cost would be willing to bid more to ensure it was able to host the research and thus get something rather than nothing.

The highest bidder assumption simply entails that researchers choose to locate their research in communities that have the lowest apparent costs (where apparent means the cost to the researcher not the actual cost of hosting the research borne by the host community). It is important to emphasize that this assumption does not entail or imply that researchers are themselves selfish, purely profit driven, or acquisitive. It implies only that they are motivated to choose host communities that reduce their costs. It may be, for instance, that researchers would like to be as generous as possible but they are under pressure from sponsors to minimize research costs. This is a plausible scenario since sponsors want to maximize the number of research studies that can be funded from a fixed research budget. In fact, this has lead some research sponsors, such as the U.S. National Institutes of Health, to limit the use of research funds to costs that are directly associated with the conduct of the research in question.

The last place assumption also appears to be fairly uncontroversial. It holds that if a community has the highest conceivable cost (i.e., c_{max}) then they expect either not to host the research, because someone will have out bid them, or to give all the surplus minus their cost to the researcher, since otherwise the researcher could have done better.

The smooth distribution assumption requires that there is some range of costs such that there are no values that are impossible. It would be strange, for instance, for a community to think that another's costs might be x or y but nothing in between.

The common knowledge assumption is perhaps the strongest since it requires both that all parties be aware of many things and also that they know what others know. Recall, however, that the proponents of the fair benefits approach suggest that eliminating informational asymmetries is a primary aim of the centralized database. The common knowledge assumption represents a sort of extreme success of this project, since the only private information is known by the individual communities. While these assumptions are required for the proof of the revenue equivalence theorem, it should not be presumed that if this constraint were relaxed more equal distributions would thus be possible. As is the case with all of these assumptions, they are sufficient, but not necessary for the proscribed outcomes.

Given these assumptions, the revenue equivalence theorem entails that regardless of the auction mechanism, the expected outcome for the researcher is the same – namely to keep all of the value minus the expected second lowest cost.

Appendix B Formal model of constrained bargaining

In this appendix we consider a non-auction mechanism which we regard as implausibly restrictive. We present it primarily as an illustration to show how even very restrictive bargaining mechanisms can produce very skewed distributions in favor of the researcher.

What if proponents of the fair benefits approach wanted to prohibit both simultaneous and sequential negotiations with multiple host communities in an attempt to equalize the relative bargaining power of the parties? Could they require, for each research project, that researchers must first choose a host community with which to partner and only then conduct negotiations about the division of benefits?

The first problem with this proposal is that it would require an implausibly strong restriction on research. In order to distinguish this from the various auction-like mechanisms, it would have to be the case that if negotiations fail the researcher is not allowed to conduct this research anywhere. Otherwise, we would have a case of multiple, sequential negotiations, and the result from auctions would apply.

Perhaps more importantly, however, even this more restrained process can still result in a race to the bottom. In one sense, this process fares better than the previous proposals, since once researchers chose a community they are committed to finding an acceptable bargain with that community and simultaneous competitive bids have been prohibited. As a result, this setup includes more equitable divisions of the surplus as possible outcomes. However, even this setup does not entirely equalize the bargaining power of the parties. The inequality persists because there will always be future studies and communities with lower costs have an incentive to lure researchers to their venues.[1] Further work is needed, therefore, in order to ascertain how plausible it is that more equitable outcomes will result from adopting this structure for negotiations.

In what follows we develop a formal model and show that in it there are a litany of potential agreements (Nash equilibria) that range from the host community getting almost the entire surplus to outcomes where the host

[1](Petryna 2007) provides an example where a host community engages in ethically questionable behavior for fear that future research by that firm would be conducted elsewhere if they adhered to more strict standards.

community expects very little surplus. If we use a more restrictive predictive tool from cooperative game theory known as the Core, we find that only a few outcomes are possible, all of which involve the host community receiving very little surplus. Informally, an outcome is in the Core if it is immune from the possibility of a coalition of players banding together to switch strategies. While there are valid criticisms of the Core we argue that there are features of this circumstance which make is a plausible predictive tool.

Informally our argument regarding the Core work like this. Suppose a bargain is struck between a researcher and host community that results in a close to equal division of resources. Since there are Nash equilibria with this result, it is a possible outcome of this setup. Now suppose, thanks to the public repository of past agreements, that another community sees the terms on which the present bargain was struck and finds that it could host a similar research project for less, allowing the researcher to keep more of the surplus. The coalition between the researcher and this new host community is better for both parties, and does not require the cooperation of anyone else. As a result, the equilibria with an equal division is not in the Core, and so we expect would not be a stable result of this process. In fact, since this possibility exists and can be predicted by others, we expect that initial bargains may be very biased as host countries attempt to prevent underbidding on future projects.

This shows, again, how this more rigid structure acts much like an auction when implemented. The only agreements that are in the Core are divisions where the host community agrees to keep very little of the surplus, if it can afford to. The amount it offers to keep must be so small, that it would not pay to switch to another community since they would only rarely be able to afford a bargain that is better for the researcher. This represents a standard race to the bottom where each community tries to eek out some profit by underbidding the previous winner. We expect that the fact that researchers engage in several different projects over time will lead to the only equilibrium that is immune from this process – one where the host community expects very little surplus.

It is this informal dynamic process that we think would be encouraged by the presence of the database. The publicly accessible record of previous negotiations actually enhances the ability of lower-cost communities to realize that they may be able to secure a benefit by underbidding current host communities. In this sense, if the database does anything, it works to the detriment of the host community.

B.1 Formal discussion

This model utilizes some of the assumptions used for the revenue equivalence theorem. We drop the Highest Bidder and Last place assumptions, and strengthen the assumption about risk to require that all parties are risk neutral. Some results will change if some of the parties are risk averse (willing to pay to avoid risk), but we do not expect this will radically alter the results. We do preserve the assumption that the researcher wishes to locate the research in the community which has the lowest cost to them.

Suppose there is a set of players $\{E, 1, \ldots, n\}$ (where E represents the researcher, and each number a community). E's strategy is a community, $i \in \{1, \ldots, n\}$ and a bottom line b_E. b_E represents the point at which E will refuse to conduct the research rather than accept a split of the value less than that. Each individual community, i, chooses a bottom line as well b_i.

Like the auction case we will assume that there is cost for the research for each individual community c_i which is drawn from some common distribution. We assume the community, but not the researcher, is aware of the actual c_i, but the distribution is common knowledge.

Let $f_i(x)$ represent the outcome of the bargaining process when the bottom lines differ by x. This represents the amount of the surplus value is allocated to the researcher when the bottom lines are compatible. As an example, suppose that the value of the research is 10. Suppose the researcher chooses community x and a bottom line of 3. Suppose community x chooses a bottom line of 5. We know that they will reach an agreement, since they have compatible bottom lines. $f_x(2)$ represents how much of the extra 2 units is kept by the researcher. So in total, the researcher receives $3 + f_x(2)$ and the community receives $5 + (2 - f_x(2)) - c_x$ (their bottom line, plus their part of the surplus, minus the cost they have to expend to support the research).

Let E's strategy be $< i, b_e >$ and community i's strategy be b_i. The players utility functions are:

$$\pi_E = \begin{cases} b_e + f_i(v - b_e - b_i) & \text{If } b_e + b_i \leq v \\ 0 & \text{Otherwise} \end{cases} \tag{1}$$

$$\pi_i = \begin{cases} b_i + (v - f_i(v - b_e - b_i)) - c_i & \text{If } b_e + b_i \leq v \\ 0 & \text{Otherwise} \end{cases} \tag{2}$$

$$\pi_j = 0 \text{ For all } j \neq i \tag{3}$$

Suppose that the $f_i(\cdot)$ functions are unknown to any players, but drawn from a commonly known distribution. Let $\mathcal{E}_f(x)$ be the expected value of $f_i(x)$ given the common distribution for f. Let $\mathcal{E}_c$ be the expected value of c given the common distribution. Then the expected utilities of the given players is given by the following equations:

$$u_E = \begin{cases} b_e + \mathcal{E}_f(v - b_e - b_i) & \text{If } b_e + b_i \leq v \\ 0 & \text{Otherwise} \end{cases} \quad (4)$$

$$u_i = \begin{cases} b_i + (v - \mathcal{E}_f(v - b_e - b_i)) - c_i & \text{If } b_e + b_i \leq v \\ 0 & \text{Otherwise} \end{cases} \quad (5)$$

$$u_j = 0 \text{ For all } j \neq i \quad (6)$$

Proposition 1 (Negotiation irrelevance) *Suppose a Bayes Nash equilibrium where E's strategy is $< i, b_e^* >$ and i's strategy is b_i^* and $b_e^* + b_i^* \leq v$, then $b_e^* + b_i^* = v$.*

Proof We will show that there cannot be a Bayes Nash equilibrium where $b_e^* + b_i^* < v$. Suppose a strategy for E, $< i, b_e >$ and a strategy for i, b_i such that $b_e + b_i < v$. Suppose that $\mathcal{E}_f(v - b_i' - b_i) = 0$ and choose a b_e' such that $b_i > b_e' > b_e$. The b_e' performs strictly better. A symmetric argument can be made for i when $\mathcal{E}_f > 0$. □

This proposition shows that, in equilibrium, the process of negotiation is irrelevant. One player or another expects to lose something in the negotiation and so has an incentive to increase their bottom line so as to not be taken advantage of in the process of negotiation. This allows us to ignore the negotiation process in considering equilibria.

Given this, we can easily see that no community will adopt a bottom line lower than their cost, since they are sure to receive a negative payoff. However, there are many different equilibria.

Proposition 2 (Any division) *Let $x = \max_e p(v-e)e$, let y be any number such that $0 \leq y \leq x$, there is a Bayes Nash equilibrium where E's strategy is $< i, y >$ and i proposes $v - y$ if $v - y \geq c_i$.*

Proof Suppose y as above. We will construct Bayes Nash equilibrium where E proposes y. Let E's strategy be $< i, y >$, i's strategy be $v - y$ if $v - y \geq c_i$ and c_i otherwise. For all $j \neq i$ let $b_j = v$. Since i never proposes less than c_i her payoff cannot be less than zero. Larger values of b_i result in zero payoff regardless of c_i, so i's strategy represents a best response for i. No alternative strategies for $j \neq i$ will increase their payoff.

Now consider the researcher. If $c_i \leq v - y$ then the researcher receives y, otherwise 0. The probability that $c_i \leq v - y$ is $p(v - y)$ and the expected utility for the researcher is $p(v-y)y$. Bidding more will result in an expected payoff of 0, and so cannot be superior. Consider $z < y$. This has expected payoff $p(v - z)z$. But the function $f(a) = p(v - a)a$ is strictly decreasing as a moves away from x. So, $p(v - z)z$ is strictly lower than $p(v - y)y$. $\square$

Essentially x in the proof represents the optimal amount to ask for when the community bids to keep c_i. When $y = x$ this is the favored equilibrium for the researcher and the most disfavored equilibrium for the community. This proposition shows that there is an improvement in this model over the previous auction model. Here there are equilibria where there are equitable divisions between researcher and host community. But this is far from the result desired. While equitable divisions are equilibria, so too are inequitable ones. So equitable outcomes are far from guaranteed. Even worse, inequitable equilibria have features which make us expect them more often than equitable ones as well – unfair equilibria are the only equilibria in the Core.

The Core represents a type of equilibria which is resilient to coalition formation. If an equilibria is in the Core, players cannot form coalitions of two or more players to move to another outcome. In this game, the only equilibria that are in the Core are those that result in the researcher keeping significant amounts of the value of the research for himself and the sponsor.

Proposition 3 *Let $y < \max_e p(v - e)e$, let E's strategy be $< i, y >$ and i's strategy be $v - y$ if $v - y \geq c_i$. Suppose all other communities strategies are such that this constitutes a Nash equilibrium. This strategy is not in the Core.*

Proof E's expected utility from this strategy is $p(v - y)y$ The payoff to all communities $j \neq i$ is 0. Consider the strategy set where E plays $< j, z >$ for some $z > y$ and $z < \max_e p(v - e)e$, and j plays $v - z$ if $v - z \geq c_j$. E's expected utility from this strategy is higher since it is closer to his preferred equilibrium (the maximum of $p(v-e)e$) and this function is strictly increasing as one approaches the maximum. This equilibrium is also better for j – his expected utility is non-zero since $z < \max_e p(v - e)e$. As a result E and j form a collation which strictly prefers this strategy set to the other. $\square$

It should be easy to see that the only strategy sets which are in the Core are those where the host community receives very low expected utility (the favored equilibria of the researcher). This result does depend on a slightly odd structure for the choice of strategy. Here the community is promising to choose a contingent strategy which is better for the researcher than the status quo strategy. However, this contingent strategy presumes that the community was unaware of its cost when making that promise. If we were to suppose that the communities where aware of their costs when offering to form the coalition with the researcher, than the results would be identical to the auction.

While the Core is occasionally criticized as an inappropriate predictive tool, we believe it is appropriate in this case. First, the non-Core equilibria are destabilized by a very small coalition. It does not require significant coordination. Second, the repeated process introduced by multiple projects over time makes the possibility of renegotiation very real. In fact, the presence of the database is likely to speed up this race to the bottom (in this case the Core), since a potential "outbidder" can use the database to determine what offer to make.